

音情報のMultiple description 符号化に関する研究

著者	魏 浩石
号	16
学位授与機関	Tohoku University
学位授与番号	情博第492号
URL	http://hdl.handle.net/10097/59943

氏名（本籍地）	WEY HO SEOK 魏 浩石
学位の種類	博士（情報科学）
学位記番号	情 博 第 492 号
学位授与年月日	平成22年 9月 9日
学位授与の要件	学位規則第4条第1項該当
研究科、専攻	東北大学大学院情報科学研究科（博士課程）システム情報科学専攻
学位論文題目	音情報の Multiple description 符号化に関する研究
論文審査委員	（主査）東北大学教授 鈴木 陽一 東北大学教授 曾根 秀昭 東北大学教授 伊藤 彰則 東北大学准教授 岩谷 幸雄 （工学研究科）

論文内容の要旨

第1章 序論

高速インターネットの普及と電子計算機の急速な性能向上を背景に、マルチメディア情報サービスの普及が急激に進んでいる。今後予想される展望として、いずれはラジオやテレビといった放送メディアとネットワーク通信との融合が考えられている。そのため、音情報や動画像情報といったマルチメディア情報を高品位に一斉配信可能とする放送型ストリーミング配信のための基盤技術の重要性は、今後一層高まると考えられる。本研究は、IPネットワークによる高品質音情報の放送型配信サービス実現において、基本的にその予防が難しく、受信側における実時間再生品質に大きく影響するパケットロスを取り上げ、その発生に対して頑健な実時間再生音質を供給可能とする高品位パケットロス補償（Packet loss concealment, PLC）処理について検討する。中でも本研究では、高品位なPLC処理の実現のため、Multiple Description (MD) 符号化法に着目する。図 1に基本的なMD符号化処理構造を示す。MD符号化とは、情報源（図1の $x(n)$ ）の冗長性を利用する、あるいは冗長性を付加することによって、情報源を元の情報量より小さく、互いに高い相関を持つ複数の部分情報（Description, 図1

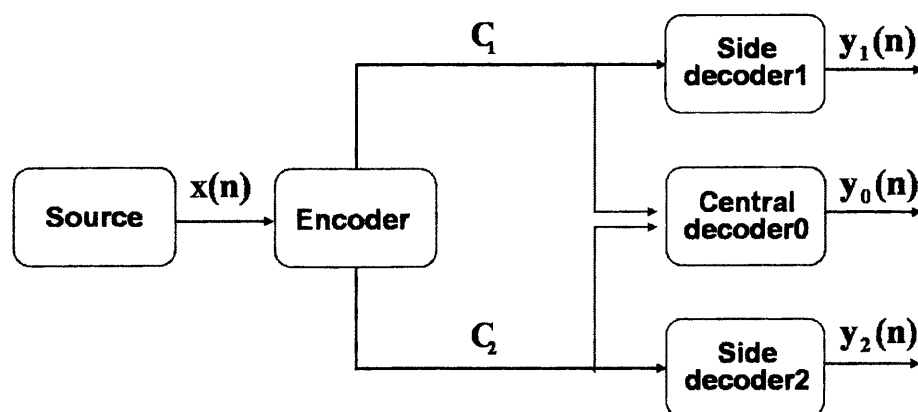


図 1: 基本的な Multiple description 符号化処理構造。

の C_1 または C_2)に分けたうえで、それぞれを通信する符号化処理のことを指す。受信側では、分けて送られてきた部分情報が全て受け取れた場合には、MD主復号器(図1のCentral decoder0)を用いて元の情報源を完全、又はそれに近い状態に復元(図1の $y_0(n)$)を行うことができる。もし通信路上での情報欠落によって一部の部分情報が欠けていたならば、受信できた残りの部分情報からMD補助復号器(図1のSide decoder1または2)を用いて一定の劣化品質で情報源を復元(図1の $y_1(n)$ または $y_2(n)$)することとなる。これにより、通信路上で情報欠落が起きても、受信側で得られる情報の品質が急激に落ちないようにすることが可能となる。つまり情報の逐次受信および復号処理が行われる際、途中で情報欠落が起きた場合に再送を行わず、補償処理のみで受信品質の大きな劣化発生を抑えることができる。そのためMD符号化は、音情報や動画像情報の実時間放送型配信分野への応用に有効とされる。

本研究では、高い必要性がありながら、従来検討があまり進んでこなかった音情報を対象とし、音情報へ効率的に適用可能な新たなMD符号化方式の開発を研究目標とする。現状の音情報通信では伝送効率を高めることを目的とする各種コーデックの導入が前提となっており、音声通信や楽曲配信といった分野において標準的なコーデックへ直接適用可能なMD符号化技術の必要性は高まっている。そのため、適用対象となるコーデックの特徴(音声通信におけるLogPCM符号化、ADPCM符号化、楽曲配信におけるMP3符号化)に合わせてMD符号化処理構造を設計することによって、既存の標準的なコーデックの修正を伴わずにMD符号化法を適用する方法の確立を目指す。

第2章 VoIP 通信のための MD 符号化技術の開発

Voice over IP (VoIP) における標準的な音声用コーデックへ直接適用可能なパケットロス補償 (PLC) 処理として、LogPCM 及び ADPCM 符号のための MD 符号化法の開発を行った。LogPCM および ADPCM 符号はスカラー量子化方式をベースとすることから、MD 符号化の概念に基づいたスカラー量子化方式である MD スカラー量子化処理を LogPCM 及び ADPCM 符号へ適用する方法について提案した。MD スカラー量子化処理は二つのスカラー再量子化テーブルを組み合わせ設計されるインデックス配置行列を用いて原情報に冗長性を付加し、二つの部分情報へ分割を行う MD 符号化法である。

LogPCM 符号に対する MD スカラー量子化の適用法として、8 ビットの符号長を持つ LogPCM 符号を、2 つの 6 ビット部分符号へ分割し、かつ振幅値が小さい再量子化区間には元の LogPCM 符号を多めに割り当てるインデックス配置行列の設計法を提案した。また ADPCM 符号に対する適用法として、4 ビット ADPCM 符号を 2 つの 3 ビット部分符号へ分割し、かつ ADPCM 符号化処理の中で量子化幅が縮小する符号と拡大する符号が再量子化区間に混在しないよう ADPCM 符号を割り当てるインデックス配置行列の設計法を提案した。これらの提案法について、パケットロスによって片方の部分符号が失われた場合の音質評価を行った。その結果、LogPCM 符号への適用結果では、部分符号のみの復号信号からでも許容範囲の音質が得られることが示された。また、ADPCM 符号への適用結果では、通常のインデックス配置行列設計法を用いて分割された ADPCM 符号の部分符号の復号信号に比べて、提案法を用いることで大幅に復号音質を改善できることが分かった。これらの結果から、提案法を PLC 処理に用いることで、パケットロス発生に対して頑健な VoIP 通信の実現が見込まれるようになった。

第3章 実時間放送型楽曲配信のための MD 符号化技術の開発

広帯域音情報の放送型配信における高品位な PLC 処理の実現を目的として、楽曲配信などの用途で

広く用いられている高品位非可逆圧縮技術である MP3 を基盤とした MD 符号化法の開発を行った。広帯域音情報のための MD 符号化法として、時間領域上の偶・奇数サンプル分割処理に基づいた MD 符号化の基本原理を MP3 コーデックと組み合わせて MD 符号化を実現する先行技術（以下、先行法）が存在する。先行法では分割された部分情報による復号音質が比較的、高音質である特徴を持つ半面、全ての部分情報を集められた場合に得られる復号音質の劣化が大きくなる問題点を持つ。本研究では、この復号音質劣化の主な要因が復号信号の 12 kHz 以上の帯域に生じるスペクトル歪みにあること、そしてその発生原因が MP3 圧縮によって生じた部分情報ごとの量子化誤差に起因することを明らかにした。そしてこの調査結果から、12 kHz 以上のスペクトル歪みを補正する Wiener filter に基づいた音質補正フィルタを設計し、主復号器の後信号処理として組み合わせて高域補正 MD 符号化法を提案した。先行法に対する提案法の音質改善効果および PLC 処理としての性能評価を行った結果、パケットロス無し/有り条件に関わらず全てのビットレート条件下で客観音質評価指標である Perceived Evaluation of Audio Quality (PEAQ) による Objective difference grade (ODG) 値の改善がみられた。また 5～10 %の高いパケットロス条件および合計ビットレート 128～192 kbps までの条件下で、他の比較手法に比べて最もよい ODG 値が得られることが確認できた。しかし、その一方で提案法による主復号信号のスペクトルには、10～12 kHz の帯域にスペクトル欠落が生じる問題は未だ残っているため、この問題は次章にて改善法の検討を行うことにした。

第4章 スペクトル補完による MD 復号情報の高品質化

前章にて提案した高域補正 MD 符号化法による主復号信号にて生じるスペクトル欠落の問題を補うためのスペクトル補完処理について検討を行った。

スペクトル補完処理を提案するにあたり、2種類のスペクトル補完成分の生成法および補完帯域成分に対する3種類のスペクトル概形生成法の組み合わせによって区分した計6種類のスペクトル補間処理について先述の PEAQ による ODG 値の音質比較を行い、最もよい補完処理方法について検証を行った。その結果、次のようなスペクトル補完法を提案した。提案法では、まず、補完対象となる主復号信号 10～12 kHz に生じるスペクトル欠落帯域に対し、その近傍である 8～10 kHz の帯域成分をヒルベルト変換による周波数シフト処理によって補完帯域成分を複写する。次に、欠落帯域の前後の平均帯域レベルを、線形補間することによって得られる推定スペクトル包絡を用いて、スペクトル概形を整形する。そして最後に、スペクトル概形が整形された補完成分の平均帯域レベルを -5 dB 減衰した上で、スペクトル補完処理を行う手順を踏む。

このような提案スペクトル補間法による効果を調べるため、第3章で提案した高域補正 MD 符号化法と本章で提案したスペクトル補完処理を併用した場合の音質改善効果について調べた。その結果、ほとんどのビットレート条件およびパケットロス率にて、音質の改善がみられた。続いて、この併用法の PLC 処理としての性能評価を行った。その結果、5～10 %の高いパケットロス条件および合計ビットレート 128～192 kbps までの条件下で、他の比較手法に比べて統計的に有意に高い ODG 値が得られることが確認できた。その一例として、合計ビットレート率が 160 kbps 条件下で平均パケットロス率が 0 % から 10 % まで増加した場合の平均 ODG 値の推移を示した結果を図 2 に示す。図 2 に示されている比較法はそれぞれ、指定された合計ビットレートで単に MP3 圧縮を行う方法 (Single MP3 条件) と、指定されている合計ビットレートの半分のビットレートで MP3 圧縮を行った音源を 2 つ用意し、同時送信する方法 (Double MP3 条件)、そして先行法 (Basic MDC 条件) と提案法 (Proposed MDC 条件) の 4 つである。この例から、提案法 (Proposed MDC 条件) は、パケットロスが 0 % 条件下では、Single

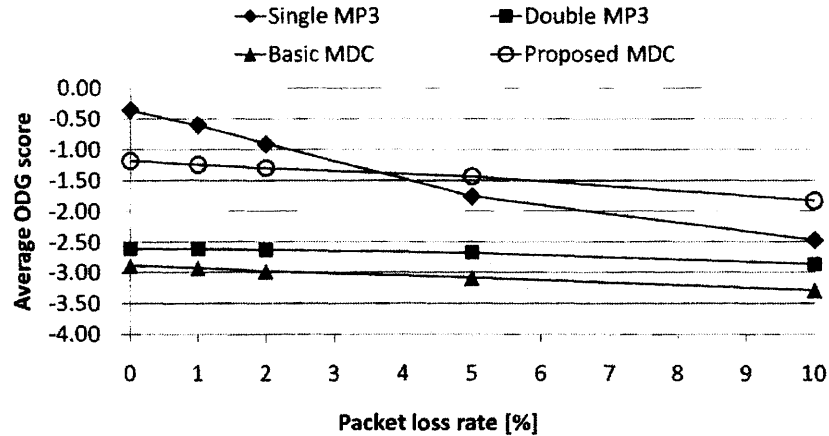


図 2: 合計ビットレート 160 kbps 条件における, ランダムパケットロス率に対する適用符号化法化法ごとの平均 ODG 値の比較結果 (合計ビットレート 160 kbps 条件の場合)

MP3条件による結果に次いで, 高いODG値が得られることがみてとれる. しかしパケットロス率が5 %以上に増える条件下では, 他の比較手法に比べてODG値が大きく劣化せず, かつ最も高いODG値が得られることが確認できる. このように提案法 (Proposed MDC条件) は, パケットロスに対して頑健に高音質が得られる様子がみられる.

この評価実験の結果から, 本章で提案した信号処理法は, パケットロス率が劇的に変動する通信路環境下における楽曲といった広帯域音情報のストリーミング配信などの用途において, 実用的かつ高音質なパケットロス補償法として利用できることが期待される.

第5章 結論

本研究では, 音情報の MD 符号化技術の開発を目標として, 音声通信において標準的なコーデックである LogPCM および ADPCM のための MD 符号化技術として, LogPCM および ADPCM のための MD スカラー量子化適用法について提案し, 提案法によって得られた部分情報のみによる復号音質からでも許容範囲内の音質が得られることが確認できた. また, 楽曲配信において標準的なコーデックである MP3 のための MD 符号化技術として, 全ての部分情報を集めた場合の復号音質に大きな音質落下が生じる問題を改善する音質補正フィルタを適用することで復号音質を改善する高域補正 MD 符号化法を提案した. 更に, 高域補正 MD 符号化法における復号信号で生じたスペクトル欠落を付加情報無しで補完する方法を提案し, これらを組み合わせた提案 MD 符号化法を用いて楽曲信号を伝送するシミュレーションを行った. その結果, 提案 MD 符号化法は, 楽曲配信において 128 から 192 kbps までの実用的なビットレートの範囲において, 5%以上の平均パケットロス発生条件下でも良好な音質を維持しうる符号化法であることが確認できた.

以上の結果から, 本論文で挙げていた目標であるネットワークを介して, かつパケットロスに頑健な音情報通信の実現が, かなりの程度達成できたものと考ええる.

論文審査結果の要旨

コンピュータとネットワーク技術の進展と社会への普及を背景に、メディア情報の実時間放送型配信が急速な広がりを見せている。しかしインターネットにおける通信ではパケットロスが不可避であることから、その影響を軽減する技術が強く求められている。著者は、この問題に着目し、音信号を対象としたパケットロス補償技術の開発を行ってきた。本論文は、その成果をまとめたもので、全編5章からなる。

第1章は序論である。ある情報源をほぼ同等の復号特性をもつ互いに冗長な複数の符号に分割して送ることにより通信の頑健性を確保する multiple description (MD) 符号化が、パケットロス補償技術として有望であることを論じている。

第2章では、音声のための MD 符号化について検討している。IP 電話に広く用いられている LogPCM と ADPCM 符号に対する MD スカラー量子化法を提案し、パケットロス時にも比較的良好な音質が確保できることを示した。これは、IP 電話におけるパケットロス耐性の向上に有効な技術である。

第3章では、楽曲等の広帯域音情報の放送型配信を目的とし、MP3 音信号に対する MD 符号化について検討している。MP3 音信号用の MD 符号化技術としては、時間領域でのサンプル分割が有望である。しかし、その先行技術では、全ての部分情報が受信できた際の復号信号の音質劣化が問題となっていた。そこで著者は、音質劣化の原因が高周波数帯域における折り返し歪であることに着目し、復号信号の後処理として Wiener filter を用いる新しい MD 符号化技術を提案した。この提案法について、伝送シミュレーションを行い、従来法との音質比較評価を行った。その結果、音質からみたパケットロス耐性が比較した従来法をしのぐことが示された。これは、広帯域音情報の実用的なパケットロス補償法として評価できる。

第4章では、第3章で提案した MD 符号の復号において、10~12 kHz 帯域の周波数成分が欠落するという問題の補正法について検討している。隣接する 8~10kHz のスペクトル成分を周波数シフトし、スペクトル包絡を調整した上で補完に用いることにより、わずかではあるが音質が更に改善された。これは第3章で提案した方法の改良法として評価できる。

第5章は、結論である。

以上要するに本論文は、パケットロス耐性に優れた音信号符号化法を提案し、その有効性を明らかにしたもので、音情報工学およびシステム情報科学の進展に寄与するところが少なくない。

よって、本論文は、博士(情報科学)の学位論文として合格と認める。